DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

AN EXPERIMENTAL STUDY OF THE EFFECT OF PRIVATE INFORMATION IN THE COASE THEOREM

Richard D. McKelvey
California Institute of Technology

Talbot Page
Brown University

SOCIAL SCIENCE WORKING PAPER 1018

September 1997

An Experimental Study of the Effect of Private Information in the Coase Theorem*

Richard D. McKelvey Talbot Page

Abstract

This paper investigates, in an experimental setting, the effect of private information on the Coase theorem's predictions of efficiency and allocative neutrality. For a two-person bargaining game, we find significantly more inefficiency and allocative asymmetry in the case of private information compared with the case of complete information. We also find substantial bargaining breakdown, which is not predicted by the Coase theorem. For the case of private information, the Coase theorem does not predict as well as a generalized version of the Myerson-Satterthwaite theorem, which predicts inefficiency, allocative non-neutrality in the direction of the disagreement point, and some bargaining breakdown.

Keywords: Coase theorem, Myerson-Satterthwaite theorem, two-person bargaining, private information, incomplete information, bargaining breakdown, cooperative and non-cooperative game theory.

*The financial support of the National Science Foundation (Grant #SBR-9223701) is gratefully acknowledged. We thank Eugene Grayver for writing the computer programs to run the experiments, and we thank Tony Lancaster and Pedro Gozalo for helpful comments.

1 Introduction

In this paper we experimentally study the effect of private (incomplete) information on the Coase theorem by making two comparisons. In the first, we compare the predictive performance of the Coase theorem with and without private information. In this comparison we find, for a specific two-person bargaining game, that the Coase theorem's predictive performance falls off substantially when we move from complete to private information.[1] In the second comparison, for the case of private information, we compare the predictive performance of the Coase theorem with the predictive performance of a generalized version of the Myerson-Satterthwaite theorem.[2] In this second comparison, we find that the generalized Myerson-Satterthwaite theorem predicts better than the Coase theorem.

As it is usually stated, the Coase theorem predicts (A) an efficient amount of harm (efficiency), and (B) an amount of harm which is the same no matter which way rights are initially assigned (allocative neutrality). An immediate consequence of (A) is the third prediction, (C) no bargaining breakdown. The generalized Myerson-Satterthwaite theorem makes three contrasting predictions: (A') inefficiency (B') systematic variation in the amount of harm in favor of the rights holder (allocative non-neutrality, or asymmetry); and (C') some bargaining breakdown.

There have been few empirical or experimental studies of the Coase theorem, assessing its predictive performance with and without private information, and to our knowledge none comparing the predictive performance of the Coase theorem with that of the Myerson-Satterthwaite theorem. In one of the earliest and most interesting ex-

[1]Coase apparently intended his theorem to be applied to environments with private information. He stated his theorem without restriction to complete information, and his applications are for cases of private information. In his example of the rancher and the farmer, it appears that neither party knows the other's cost or value (pp. 97-102); in his examples of actual legal cases, private information is the only plausible assumption (pp. 104-113); in his argument against Samuelson's claim of bargaining indeterminacy when there is private information he argues for an efficient bargaining outcome for the same information environment (pp. 160-3); and his argument against Pigovian taxes takes place in an environment of private information, where he argues that in many cases direct bargaining will be more efficient than trying to overcome the information requirements of Pigovian taxes (pp. 182-4), [Coase, 1990; references to "The Problem of Cost" and "Notes on the Problem of Social Cost" here as well.]

[2]Fudenberg and Tirole point out the conflict, p. 279, see also p. 245.

perimental studies of the Coase theorem, Hoffman and Spitzer [1985] found very high efficiencies with and without private information, and little difference between the two information environments. This finding differs from ours, of a substantial impact of private information in lowering efficiency.[3]

Hoffman and Spitzer provide an example of policy recommendation based on the Coase theorem and their evidence in support of the theorem. They write: "The almost complete dominance of Pareto optimal outcomes in our two-person experiment suggests that, if there is only one homeowner, a court may choose between [the polluter's right and the victim's right] ...with confidence that the parties will bargain to an efficient outcome. Hence, injunctive entitlements have appeal in two-party situations" (Hoffman and Spitzer [1985], p. 97).

In contrast, the results of our experiment suggest more caution in basing policy recommendations on the Coase theorem, when there is private information. Adding to this caution, we conjecture that in cases of bargaining over actual environmental harms, where the structure of information and the payoff functions are more complicated, the amounts of harm are more than one-dimensional, and the number of bargainers more than two — that in such realistic cases the impact of private information may decrease the efficiencies below what we observed in our experiment for the incomplete information treatment.

The results of the paper can be viewed from another perspective as well: a comparison of approaches of cooperative and non-cooperative game theory. The Coase theorem shares the approach of cooperative game theory in several ways: the information struc-

[3]There are several differences between the Hoffman-Spitzer experiment and ours which may explain the difference in findings. In the Hoffman-Spitzer experiment, bargaining was face to face. The experimental subjects were explicitly allowed to reveal their private information and verify it by showing their worksheets to each other during the experiment if they so chose. Over half of the subjects chose to do so [personal communication with Matthew Spitzer (6/6/97)]. In our experiments, to control information flows and keep the complete and incomplete information treatments separate, communication was done through computers. In the Hoffman-Spitzer designs, there were seven or eight possible choices for the amount of harm ("activity level"); in ours there were 101 possible choices for the amount. Thus, the task of finding the efficient level, when information is private, was probably more difficult in our design. Finally, in the Hoffman and Spitzer experiments, there was unlimited time for bargaining, while we imposed time limits.

ture and the institutional process of reaching a decision are both unmodeled. Instead, the characteristics of the solution concepts are built in "axiomatically" or by assumption.[4] As we shall see, for the bargaining game studied in this paper, the core, the Nash bargaining solution, and the von Neumann-Morgenstern bargaining set make the same three predictions made by the Coase theorem. And as with the Coase theorem, while cooperative game theory has often been developed with tacit or explicit assumptions of complete information, its solution concepts have sometimes been applied to situations with private information. For example, the First Welfare Theorem is often stated: "The competitive equilibrium is in the core," and this implication is often used to predict that market processes will tend toward efficiency, even when some information is private.

The obvious question to ask of this approach in cooperative game theory is how well do its predictions perform when there is private information, the important case for policy development. This question is sharpened by the contrasts between approaches and predictions in non-cooperative game theory.

In non-cooperative game theory the information structure and the institutional process are explicitly modeled. The information structure is precisely specified and the institutional process is explicitly defined in the extensive form of the game (or by specific rules of the game, as we do in the experiment). The solution concepts, such as a Bayesian equilibrium, often, but not always, imply inefficiency. But solution concepts in non-cooperative game theory also often imply multiple equilibria; there may be coordination problems in selecting an equilibrium; and often discovering an equilibrium requires complicated analysis. The same question arises: how well does the approach in non-cooperative game theory predict for actual decision problems?

While the differences in approach between cooperative and non-cooperative game theory can be clarified by theoretical analysis, empirical investigation is needed to assess predictive performance. It is difficult to estimate actual efficiencies econometrically, when the underlying values and costs are private information. But in experiments, the

[4]William Samuelson (1984) has pointed out the parallel between the Coase theorem and cooperative game theory.

experimenter can "induce" incomplete values and costs, facilitating their estimation. Perhaps because of this advantage, there have been an increasing number of experimental studies on the predictive performance of cooperative and non-cooperative game theoretic solution concepts.

Under some experimental conditions, solution concepts from cooperative game theory have predicted very well, with and without complete information. When there is private information the core has predicted well for experimental markets, double auctions, and committee voting (Ordeshook, pp. 370-6). Also in a bargaining experiment with private information Roth and Malouf found the Nash bargaining solution predicted better in an incomplete information environment than in a complete information environment (Holt and Davis, p. 251). Non-cooperative solution concepts have also predicted well, under some experimental conditions.

The results of our paper suggest that for the specific conditions of our experiment, the non-cooperative approach as exemplified by the Myerson-Satterthwaite theorem predicts better than does the core, the Nash solution, or the von Neumann-Morgenstern bargaining set.

The paper is organized as follows. In Section 2, we construct the bargaining game of the experiment. In Section 3 we discuss two assignments of rights – polluter's rights and victim's rights – and show how both reduce to the bargaining game of the experiment. In section 4, we identify predictions for this game from Coase, cooperative game theory, and the competitive equilibrium. In Section 5, we identify contrasting predictions from a generalized version of the Myerson-Satterthwaite theorem. In Section 6, we report the results of experiments run on the bargaining game.

2 Design of the Experiment

In this section, we do three things. First, we list our design goals. Second, we specify the bargaining game used. Third, we explain how the game was operationalized in our experiments.

2.1 Design Goals

To facilitate comparison of the Coase theorem and the Myerson-Satterthwaite theorem, our main design consideration was to construct a bargaining game which fits the conditions of both theorems. More specifically, we incorporated the following design characteristics:

- Assignments of rights was defined in terms of a disagreement point, which is made known as common knowledge. This feature fits with the Myerson-Satterthwaite theorem and has been used in analyses of the Coase theorem (see Hurwicz, 1995). This design feature simplifies our comparison of the two theorems.

- Additive separable payoff functions. The Myerson-Satterthwaite theorem incorporates additive separable payoffs, and one of the sufficient conditions for the implication of allocative neutrality in the Coase theorem is additive separable utility (Hurwicz, 1995). We can't guarantee additive separable utility in an experiment, but the additive separable payoffs work in that direction.

- Free-form bargaining. We attempted to define a bargaining process which was as unstructured as possible. The Myerson-Satterthwaite theorem applies to virtually any two-person bargaining process, and in much of the literature on the Coase theorem the bargaining process is either explicitly free form (e.g., the Hoffman-Spitzer experiment) or unspecified (e.g., Coase's example of the farmer and the cattle-raiser [Coase, 1960]).

- Two-person bargaining. Much of the literature on the Coase theorem focuses on this simpler case, and two-person bargaining is a feature of the Myerson-Satterthwaite theorem.

- Strict control of information flows. We wanted to maintain a clear separation between complete and private information treatments. Private information is a condition of the Myerson-Satterthwaite theorem, and a blurring of this distinction, through face-to-face bargaining, may have been a source of the observed high efficiency in the Hoffman-Spitzer experiment.

- A divisible amount of harm. Much of the policy application is for a divisible amount of harm and most of the literature on the Coase theorem is for a divisible amount of harm. The Myerson-Satterthwaite theorem is only for an indivisible unit of a good (or harm) and to make our comparison, we derived in a theoretical paper (McKelvey and Page, 1996) a generalized version of the Myerson-Satterthwaite theorem to apply to a divisible harm.

- Low transaction costs. This is a requirement of the Coase theorem.

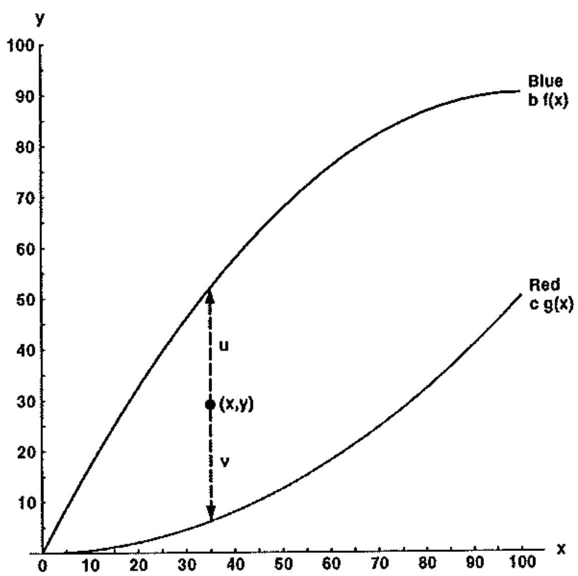
- Quadratic cost and benefit functions. Quadratic cost and/or benefit functions have been used previously in the literature on the Coase theorem. Coase, in his table and discussion of the rancher and farmer, uses a quadratic cost function for the farmer (pp. 111-116 DD); Turvey (p. 311, Figure 1), Demsetz (p. 68, Figure 1), and Mishan (p. 20, Figure 2) have quadratic cost and benefit functions of pollution.

2.2 Specification of the Game

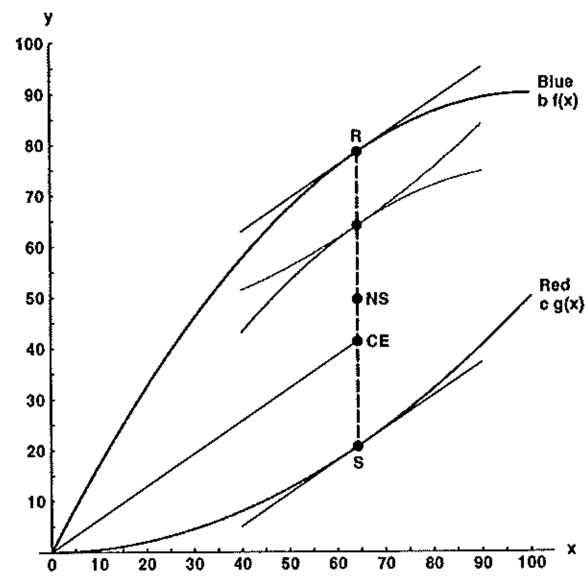
The following bargaining game incorporates the above design goals. There are two players, "Blue" and "Red", who bargain over the location of a point in a square, $A = [0,100] \times [0,100]$. Define $f, g : [0,100] \mapsto \Re$ by $f(x) = 2x - \frac{x^2}{100}$, and $g(x) = \frac{x^2}{100}$. If the two players agree on a particular point $(x, y) \in A$ then their payoffs are:

(1)	$u = bf(x) - y$	Blue player's payoff
(2)	$v = y - cg(x)$	Red player's payoff
(3)	b, c	Independent and uniform random variables, $b \in \{0.1, 0.2, \ldots, 1.0\}$ each w.p. 0.1 and $c \in \{0.1, 0.2, \ldots, 1.0\}$ each w.p. 0.1
(4)	$(x, y) = (0, 0)$	Status quo, or disagreement point

where x and y are integers between 0 and 100.



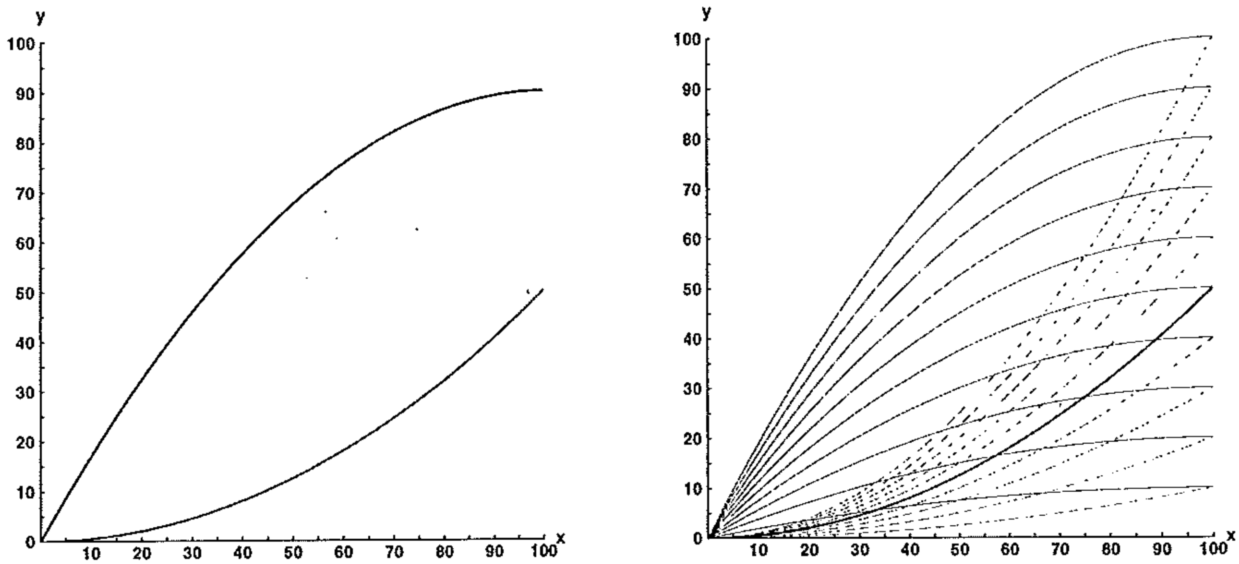
(a) Payoff Functions



(b) Predictions

Figure 1
Sample payoff functions for $b = .9$, $c = .5$
The efficient x is at $x = 100b/(b + c) = 64$

The payoff functions are illustrated in Figure 1(a) for the case where $b = .9$ and $c = .5$. In the experiment, Red and Blue must either come to an agreement on a point (x, y) in the unit square or end up with $(0, 0)$ as the disagreement point. If (x, y) is the agreed-upon point, then Red's payoff v is the vertical distance of (x, y) above Red's curve $cg(x)$; Blue's payoff, u, is the vertical distance of (x, y) below Blue's curve $bf(x)$. If there is no agreement, $(0, 0)$ is chosen, the distances are zero and $u = v = 0$.



(a) Full Information

(b) Private Information (Red's Screen)

Figure 2
Sample computer display for $b = .9$, $c = .5$

2.3 Operationalization

The above game was set up to run over a computer network, with subjects only able to communicate with each other over the computer network. We ran two experiments on separate days, using a total of 26 subjects (14 in the first experiment, and 12 in the second, all Caltech undergraduates). An experiment consisted of a number of matches. At the beginning of each experiment, the subjects were randomly assigned a color (Blue or Red) which remained the same during the entire experiment. They then participated in a sequence of bargaining sessions. In each session they were randomly matched with a subject of the opposite color. The design was to have each subject participate in four complete information sessions, followed by nine incomplete information sessions, and then finish with four complete information sessions.[5] This yielded a total of 76 $(28 + 24 + 24)$ two-person bargaining sessions with complete information, and 109 $(55 + 54)$ bargaining sessions with private information.

As mentioned above, we ran two treatments, one with complete information and one with incomplete (private) information. In both of the treatments, subjects know, as common knowledge, the payoff functions, the pdf's of b and c, the disagreement point, and the domains of x and y. In the game with private information observations on b and c are not common knowledge. After the random variables are drawn, Blue learns b but not c, and Red learns c but not b and both subjects are informed publicly of this information structure. In the complete information game, both subjects learn both b and c as common knowledge.

In the experiment, each subject privately sees his/her own computer screen, which looks much like Figure 2(a) when information is complete, and like Figure 2(b) when information is incomplete. For the complete information treatment, subjects see both their own payoff curve and the curve of the subject that they are matched with. In the private information treatment, Red sees the Red curve but not the Blue curve, and Blue sees the Blue curve but not the Red curve (both are publicly told this). The possible curves for *all* types of both subjects are displayed using dotted lines, but the subject's own actual payoff function is highlighted. For both games, both subjects are publicly told what curves each sees.

In the experiment the subjects can determine the approximate payoffs from any potential outcome $(x, y) \in A$ by moving their mouse to that point on the screen. A display on their computer screen gives their payoff from that point. This payoff is continually updated with the correct payoff as they move the mouse around the screen.

[5]In the first experiment, due to a computer malfunction, the experiment was terminated at the end of the eighth incomplete information session, and the last four complete information sessions were not conducted. We therefore only obtained 55 observations for the incomplete information treatment, and none for the final complete information treatment in that experiment.

The experiment is run in continuous time over a finite period (five minutes), with no restrictions that subjects must alternate in making offers, and with no restriction on the number of offers that can be made by either subject. If the subjects agree on a point (x, y), each gets the payoff defined above; if there is no agreement, both get zero payoff.

To make an offer, the subject clicks her mouse on a point. When an offer is made, a small cross is displayed on her screen and a small disk is displayed on the screen of the subject she is matched with. Either subject can revise an offer by making another offer. Only one offer for each subject is valid at a time. Either subject can accept an offer made by their partner by clicking on the disk token representing the other's current offer. If this occurs prior to the termination of the session, then the session ends with the agreed-upon point used to compute the payoffs of the subjects. If the subjects do not come to an agreement prior to the termination of the experiment, then the payoff is determined by the disagreement point, so neither subject gets any payment for that session.

In the experiment, the subjects are given no interpretations about their roles, the choice variables, or the payoff function. Each subject communicates through a computer network. The Red subject is just referred to as the Red subject, the function $cg(x)$ shown on Red's screen is colored red and called the Red curve, c is not explained as a "utility-relevant parameter" or "Red's type" but as part of the definition of the Red curve, x and y just describe a location on a square shown on a computer screen, etc. Subjects make offers by clicking on locations of their cursor on their computer screen, and one subject can accept the other's offer by clicking on it.

We will call the above bargaining game the "uninterpreted experiment." In the instructions to the experiment there is no mention of "rights," "polluter," "victim," "utility parameter," or to Coase or Myerson-Satterthwaite. Even the words "bargaining game" are not used, to avoid their possible connotations, and the experiment is called more neutrally, "an experiment in decision making."

3 Assignment of Rights

The above experiment can be interpreted with two differing assignments of rights.

- *Victim's rights.* Here x is the amount of an environmental harm, say smoke. The victim, identified as the Red subject, bears a cost of $cx^2/100$ from x, the level of smoke emitted by the polluter. The polluter, identified as the Blue subject, benefits $b(2x - x^2/100)$ from emitting x.

  The victim has the right to clean air. The right is alienable in the sense that the victim can sell off part or all this right, for monetary compensation y from the polluter. If the victim and polluter can reach a mutual agreement on the level of pollution x and the payment y from the polluter to the victim, then (x, y) is the outcome to the bargaining process. If the victim and polluter fail to reach a mutually voluntary agreement, the outcome of the process is no pollution, no compensation, and $(x, y) = (0, 0)$.

  With this assignment of rights the payoffs and disagreement point are as (1), (2) and (4). Condition (3) is interpreted as specifying the subjects' types.

- *Polluter's rights.* In this second assignment of rights the polluter has the right to emit 100 units of smoke. As part of this rights entitlement the polluter also has the legal obligation to pay a lump sum tax of 100b. In this second assignment of rights, the victim has no right to clean air, but instead has the right to a lump sum subsidy of 100c. The polluter's right to emit 100 units of pollution is alienable in the sense that if the polluter and victim can agree to some level of abatement and some amount of compensation y paid by the victim to the polluter, then the remaining amount of pollution x and compensation y become the outcome of the game. Without agreement the outcome is 100 units of pollution and no compensation (the disagreement point is $(x, y) = (100, 0)$).

The two assignments of rights are quite different. Under victim's rights, disagreement leads to no pollution under polluter's rights disagreement leads to 100 units of pollution. But the second assignment of rights also reduces to the same uninterpreted game (1) - (4).

To see this we write the payoffs

$$
\begin{aligned}
u &= b(2x - \frac{x^2}{100}) + y - 100b \\
&= b(2(100-z) - \frac{(100-z)^2}{100} + y - 100b \\
&= y - \frac{bz^2}{100} \\
v &= -y - c\frac{x^2}{100} + 100c = -y - c\frac{(100-z)^2}{100} + 100c \\
&= c(2z - \frac{z^2}{100}) - y
\end{aligned}
$$

We interpret $z = 100 - x$ as the amount of abatement, $b(2x - x^2/100)$ as the polluter's benefit from x, y as the payment to the polluter from the victim, $100b$ as the tax, $cx^2/100$ as the victim's cost of pollution, and $100c$ as the subsidy.

With the interpretation of polluter's rights we identify the Red player as the polluter, the Blue player as the victim, and the amount along the horizontal axis as the amount of abatement z. The disagreement point $(x, y) = (100, 0)$ then becomes $(z, y) = (0, 0)$. To finish the interpretation we interchange the names of the random variables, b and c, (which have identical distributions). Then the interpreted bargaining game with polluter's's rights has identical structure with (1) – (4).

This equivalence means that in half of the experimental bargaining sessions we could have victim's rights, and polluter's rights in the other half, but when we actually run the experiments blind (striped of the interpretive words in pollute, "victim," "rights", etc.), we would use the same instructions and game form for both rights assignments. Because the two assignments of rights reduce to the same uninterpreted game, we are able to test the Coase theorem with the same experimental game.

The bargaining game can also be given an interpretation corresponding to the usual interpretation of the Myerson-Satterthwaite theorem. Blue is the buyer, $bf(x)$ is the benefit to the buyer from buying x amount of the good, Red is the seller, $cg(x)$ is the cost of producing and selling x, the rights are assigned so that the seller initially owns the good, and y is the payment from the buyer to the seller.

4 Predictions from Coase, Cooperative Game Theory, and the Competitive Equilibrium

The Coase theorem predicts an efficient amount of x, and this amount is easy to calculate. To find the efficient amount for a particular realization of b and c, maximize the social surplus $u+v = bf(x) - cg(x) = b(2x - x^2/100) - cx^2$, over x, getting the efficient amount of x to be $100b/(b+c)$. In Figure 1(b) the efficient x is the value of x which maximizes the gap between the Red and Blue curves. In the figure, all the points on the dotted line RS are efficient. To allow for error and Coase's prediction to be approximate rather than exact, define

$$x = 100b/(b+c) + r$$

where x is the actual amount chosen in a bargaining session with utility-relevant parameters b and c, and r is the residual difference between the actual amount of x and the efficient amount.

For Coase's prediction of efficiency to be exact we must have $r = 0$ for every b and c. For Coase's prediction of efficiency to be approximately true we must have $E(r)$ and $var(r)$ to be "small." For a Coasian prediction of approximate efficiency to extend from the case of complete information to private information with little difference in predictive accuracy we must have $E(r)$ and $var(r)$ about the same with and without complete information.

Coase's prediction of allocative neutrality can be interpreted as a claim that $E(r) = 0$ even if $var(x) > 0$. To see this consider what would happen if $E(r) \neq 0$. For example, if $E(r) < 0$, then with the interpretation of victim's rights x is the amount of pollution and $E(r) < 0$ implies that the amount of pollution is systematically less than the efficient amount. With the interpretation of polluter's rights x is the amount of abatement and $E(x) < 0$ implies that the amount of abatement is systematically less than the efficient amount, and hence, the amount of pollution is systematically more than the efficient amount of pollution. But then the amount of pollution varies systematically with the assignment of rights, contrary to Coase's claim of allocative neutrality.

We define a bargaining breakdown as occurring whenever the actual $x = 0$ but the efficient x is positive. In the experiment $b > 0$ and $c > 0$, so the efficient x is always positive. We conclude that whenever the observed $x = 0$ we have a case of bargaining breakdown. Thus an implication of Coase's prediction of efficiency is that there will be no observed breakdowns (no cases of $x = 0$). More practically, a prediction of approximate efficiency (where $E(r)$ and $var(r)$ are both small) is that there will be few cases of bargaining breakdown.[6]

[6]The prediction of no breakdowns, although an immediate consequence of the efficiency claim, is less discussed in the Coasian literature than the claim of allocative neutrality with respect to rights assignments, even though the neutrality claim is more delicate since it requires quasi-linearity. In Coase's "The Problem of Social Cost" we find eight explicit claims that the amount of harm will be the same no matter what is the initial allocation of rights, and only two claims that there will be no bargaining breakdowns, and these latter statements are indirect. On p. 114 Coase writes "such a rearrangement of rights will always take place if it would lead to an increase in the value of production." Thus he indirectly claims that the disagreement point will be avoided when it is efficient to do so, (pp. 114. In bargaining between firms Coase uses efficiency and maximum value of production synonymously. On p. 115 he restates this claim with the implication that there will never be bargaining breakdown. But later, in his "Notes on the Problem of Social Cost," while criticizing Samuelson's view that bilateral bargaining can lead to indeterminacy and inefficiency, Coase softens his implied claim of no breakdowns by writing it is "impossible to argue that two individuals negotiating an exchange must end up on the contract curve, even in a world of zero transactions costs. ... However, there is good reason to suppose that the proportion of cases in which no agreement is reached will be small." (See pp. 101, 102, 103, 104, 106, 108 and 110.)

For future reference we gather these Coasian predictions:

(A)	For all b and c, $r = 0$	(full efficiency)
	$E(r)$ and $var(r)$ are both "small"	(approximate efficiency)
(B)	$E(r) = 0$	(allocative neutrality)
(C)	For all b and c, $x > 0$	(no bargaining breakdown)
	For all b and c, $Prob\ (x = 0)$ is small	(breakdown rare).

These predictions also follow from other solution concepts which predict efficiency. These include the core, a competitive equilibrium, the von Neumann-Morgenstern bargaining set, and the Nash bargaining solution because they all agree with the Coase theorem's claim of efficiency. These solution concepts are illustrated in Figure 1(b). Assuming that the bargainers attempt to maximize their expected payoff we can interpret payoffs as utilities, and in the figure, with victim's rights, $b(2x - x^2/100)$ is not only the polluter's benefit from pollution x, but $bf(2x - x^2/100)$ is also an indifference curve for Blue, the locus of points (x, y) where $u = 0$; similarly $cx^2/100$ is Red's indifference curve where $v = 0$. With quasi-linear utility, Blue's indifference curves are vertical translations of $bf(x)$ and similarly for Red. The contract curve is a vertical line and it goes through the maximum gap between $bf(x)$ and $cg(x)$. The segment *RS* is that part of the contract curve which also satisfies individual rationality ex post (IR). Thus, for the two-person game, *RS* is the core, and the von Neumann-Morgenstern bargaining set is also *RS*. Coase also predicts bargaining will lead to an allocation which is both efficient and satisfy IR, in other words also in *RS*. The competitive equilibrium predicts the point *CE* where a line, through (0,0) and parallel to the tangent at it S, intersects *RS*. The Nash bargaining solution predicts the point *NS*, which provides an equal sharing of the maximum social surplus over the disagreement point.

5 Predictions from the Generalized Myerson-Satterthwaite Theorem

For any b, c and x the social surplus is $u+v = bf(x) - cg(x) = 2bx - (b+c)x^2/100$. Recall that the social surplus is maximized at $x = \frac{100b}{b+c}$. Thus the maximum social surplus for b and c is

$$2b\left(\frac{100b}{b+c}\right) - \frac{b+c}{100}\left(\frac{100b}{b+c}\right)^2 = \frac{100b^2}{b+c} \tag{1}$$

We define the **ex post efficiency** of a given x to be the social surplus divided by the social surplus of the efficient allocation. So the ex post efficiency of x, given b and c, is $e(x; b, c) = \frac{b+c}{100b^2}(2bx - (b+c)x^2/100)$. The **expected efficiency** of the allocation rule x is defined as $E[e(x(b,c); b, c)] = E[\frac{b+c}{100b^2}(2bx - (b+c)x^2/100)]$.

The Myerson-Satterthwaite theorem [Myerson-Satterthwaite, 1983] applies to an indivisible good. We have derived (McKelvey and Page, 1996) a generalized version of the theorem which applies to a divisible amount of a good (or harm) and to the bargaining problem of the experiment. The result is:

Proposition *Assume individual utility functions and information are as in (1) - (4). And define, as before the residual* $r = r(b,c) = x(b,c) - 100b/(b+c)$. *Then any incentive compatible direct mechanism* $(x(b,c), y(b,c))$ *satisfying individual rationality (ex.ante or ex post) also satisfies:*

(A') $E[e(x; b, c)] < 1$. (inefficiency)

(B') $E(r) < 0$ (allocative asymmetry)

There appear to be a very large number of Bayesian equilibria for the bargaining game with private information. We have calculated just a few, finding in one Bayesian equilibrium bargaining breakdowns five percent of the time, in another Bayesian equilibrium breakdowns ten percent of the time, and in a third Bayesian equilibrium breakdowns 14 percent of the time. (These three examples suggest that generalized versions of the Myerson-Satterthwaite theorem might provide a theory of strikes.) Since we know that at least some Bayesian equilibria have bargaining breakdowns, and we have no theory ruling out Bayesian equilibria with breakdowns, we make a third prediction:

(C') $\quad Prob(x = 0) > 0 \quad$ (Some bargaining breakdown)

Further analysis predicts that bargaining breakdowns are more likely to occur when b is low and c is high. Our calculations of a few Bayesian equilibria tentatively suggest the hypothesis that the frequency of breakdown will be in the range of 2-20%.

The three predictions of the Coase theorem (A), (B) and (C) can be tested in treatments with and without complete information, and in doing so the effect of private information can be studied. Here information is the treatment variable.

For the treatment with private information the conflicting predictions of (A) and (A'), (B) and (B') and (C) and (C') can be compared. Here we are comparing two different models and solution concepts in the same treatment with private information, and we can study how well the predictions of the Coase theorem apply to an environment with private information by comparing them with predictions from the generalized Myerson-Satterthwaite theorem. Each of the predictions of the generalized Myerson-Satterthwaite theorem's predictions, (A'), (B') and (C') is at odds with the Coase theorem.

6 Experimental Results

We are now ready to address the question we asked at the beginning of the paper: "How well does the Coase theorem predict when there is private information?" In our first comparison (see Table 1), we compare the predictive performance of the Coase theorem with and without complete information. Let n be the total number of observations, and let the k subscript denote the values of the corresponding variables in the k^{th} bargaining session. Average observed efficiency (defined as a percent) is defined by

$$\frac{1}{n}\sum_k (2b_k x_k - (b_k + c_k)x_k^2/100)(b_k + c_k)/b_k^2,$$

allocative neutrality is defined by $\frac{1}{n}\sum_k r_k$, and percent breakdowns is defined by $\frac{100}{n}\sum_k BR_k$, where $BR_k = 1$ if the k^{th} bargaining session ends in breakdown and 0 otherwise.

	Predicted	Observed	
		Complete Inf. (n=79)	Private Inf. (n=109)
Efficiency (%)	100	91	67
Allocative Neutrality	0	-1.7	-16.1
Breakdowns (%)	0	7.9	27.5

Table 1: Coase Theorem's Predictions Compared with Experimental Outcomes, for the Two Treatments

From Table 1 it appears that the Coase theorem does not predict as well when there is private information compared with the complete information treatment. To check whether the observed differences are statistically significant, note that almost everything depends on the distribution of the residuals r. If the mean and variance of r are both small, then efficiency will be close to 100%, allocative neutrality will be close to zero, and bargaining breakdown will be rare. If the average of observed r is small (even with a large variance), there will be approximate allocative neutrality, and there will be few breakdowns. Thus, a test of whether there is much difference in predictive performance between the two information treatments is to test for differences between means and variances of r, in the two treatments.

Write $var_c(r)$ for the variance of r in the complete information treatment; $var_p(r)$ for the variance of r in the private information treatment; $E_c(r)$ for the mean of r (expectation over b and c) in the complete information treatment; and $E_p(r)$ for the mean of r in the private information treatment. Using an F test for

$$H_0: \quad var_c(r) = var_p(r)$$

$$H_1: \quad var_c(r)/var_p(r) \neq 1$$

We find $F = 0.34$, p-value < 0.0001, num $df = 75$ den $df = 108$, sample $var_c(r) = 116$, sample $var_p(r) = 347$.

Using a Welch modified two-sample t-test for

$$H_0: \ E_c(r) = E_p(r)$$
$$H_1: \ E_c(r) \neq E_p(r)$$

we find $f = 6.6$, p-value < 0.0001, $df = 177.7$.

The two tests confirm that the effect of private information on the Coase theorem is statistically significant.

In our second comparison, for the case of private information, we compare the predictive performance of the Coase theorem with that of the generalized Myerson-Satterthwaite theorem, Table 2. For each of the three predictions, the Myerson-Satterthwaite theorem comes closer to the observed outcomes than does the Coase theorem. Again, everything depends on the distribution of r. A simple test of the Coase theorem's prediction of allocative neutrality compared against the generalized Myerson-Satterthwaite theorem's prediction of allocative asymmetry toward the disagreement point is:

$$H_0: \ E_p(r) = 0 \quad \text{(Coase prediction of allocative neutrality)}$$
$$H_1: \ E_p(r) < 0 \quad \begin{array}{l}\text{(Myerson-Satterthwaite prediction} \\ \text{of allocative asymmetry)}\end{array}$$

For this we used a one-tail t-test; $t = -9.0$, $df = 108$, $p - value < .0001$. We reject the Coasian hypothesis in favor of the Myerson-Satterthwaite alternative.

A more detailed inspection of the sequences of offers and counter offers suggests that the bargaining process, when there is complete information differs substantially from the process with private information. In nearly all of the bargaining sessions with complete information, the offers and counter-offers went quickly to the efficient level of x

	Predicted		Observed
	Coase	M-S	(n=109)
Efficiency %	100%	$< 80\%$	67%
Allocative Neutrality	0	< 0	-16.1
Bargaining Breakdown %	0	$\sim$ 4-14	27.5

Table 2: Predictions from the Coase Theorem Compared with Predictions from the Myerson-Satterthwaite Theorem for the Case of Private Information

and then bargaining continued over the transfer payment y (in the complete information treatments the bargaining looked a lot like a divide-the-dollar game). In many cases an equal or close-to-equal split was offered and then accepted (with efficiency close to 100%). But occasionally one subject (or both subjects) would hold out for a larger share of the maximum social surplus, and sometimes this would lead to bargaining breakdown. It appears that most of the inefficiency in the complete information treatments comes from this type of bargaining breakdown (in which a symmetric game of divide the dollar becomes a game of chicken). It appears that even with occasional bargaining breakdown, the Nash bargaining solution predicted well, when there was complete information, because close to equal splits of the maximum social surplus were frequently observed.

The process of offers and counter offers looks different in the treatments with private information. Here the game does not look like a divide-the-dollar game. Not only do the subjects not know where the maximum social surplus is, they even have considerable trouble discovering values of x which satisfy individual rationality for both subjects (it appears that some of the breakdowns arise because they fail to find, with any confidence, values of x which are mutually individually rational, especially when b is low and c high).

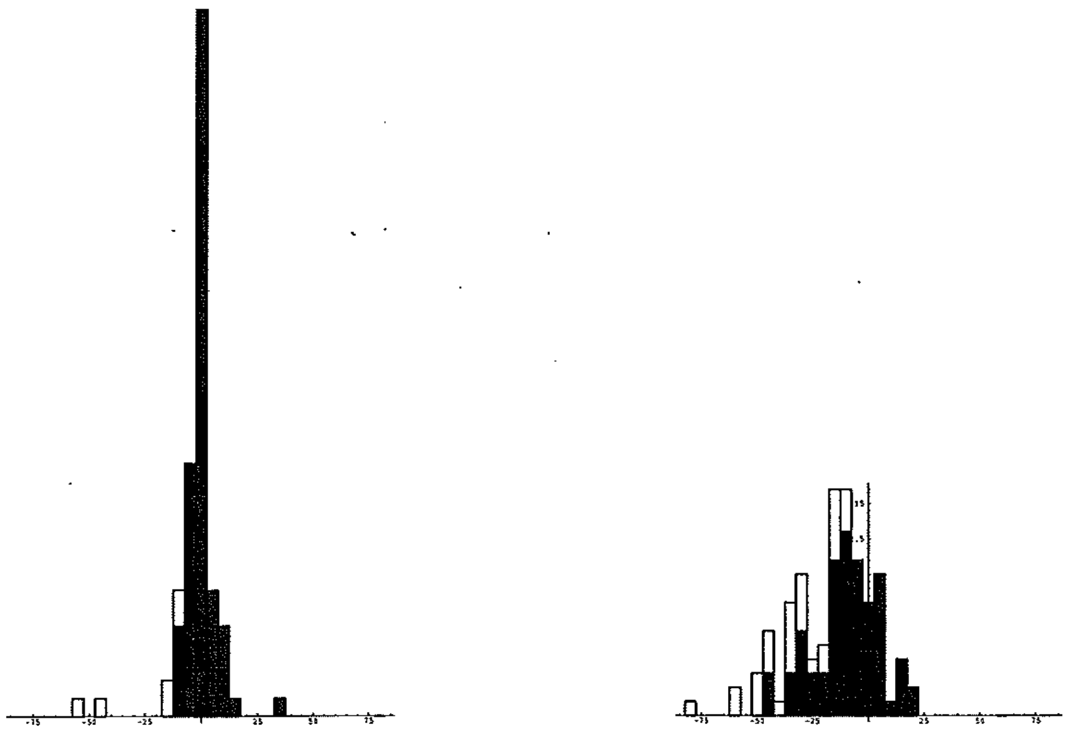
In treatments with private information the observed pattern of bargaining is consistent with the usual interpretation of bargaining over a divisible good — the buyer systematically understates his value and begins by asking for large amounts of x at low prices; the seller systematically overstates her cost and begins by offering large amounts of x at high prices. Gradually both subjects soften their demands. The buyer raises his bid price and also decreases the amount of x offered to be bought. Correspondingly, the seller lowers the asking price and also lowers the amount of x offered to be sold. This systematic lowering of the x offers is sensible, since the subjects know as common knowledge that the marginal benefits of the good to the buyer increase as x decreases and the marginal costs to the seller decrease as x decreases.

There are many fewer equal splits when there is private information, which is not surprising because the subjects have no way of knowing what would be an equal split. Thus, it appears that bargaining with private information follows a quite different dynamic; the pattern for both subjects of initially high and then falling offers of x is not observed when there is complete information.

Figure 3 shows histograms of the residuals r for the two treatments. In the treatment of complete information (Figure 3(a)) the deviations are roughly symmetric around zero, as predicted by (A) and (B) of the Coase theorem.

But for the treatment with private information (Figure 3(b)), the residuals r tend to be negative, as predicted by the Myerson-Satterthwaite theorem and contrary to the Coase theorem. The asymmetry is substantial, and can be roughly interpreted as follows: with an average of the negative r's on the order of 20-25 units of x, with polluter's rights when the efficient level of abatement is 50 the observed level of abatement could easily be 30; with victim's rights when the efficient level of pollution is 50 the observed level of pollution could easily be 30. In other words, with an average negative divergence of 21-25, we could easily see 70 units of pollution with polluter's rights and only 30 with victim's rights — a result strongly at odds with Coase's claim that the allocation of rights has no effect on the level of pollution.

In Figure 3 the cases of bargaining breakdown are shown as unshaded. Clearly the observed inefficiency and allocative asymmetry occurs even when the cases of bargaining breakdown are excluded from the analysis. Bargaining breakdowns are anomalies for the Coase theorem, but predicted by the Myerson-Satterthwaite theorem.



(a) Complete Information (b) Incomplete Information

Figure 3
Residuals in complete vs. incomplete information experiments
(Unshaded portion results from bargaining breakdown)

In summary for the experiment: The Coase theorem for complete information predicts pretty well but not spectacularly for efficiency (A), reasonably well for allocative neutrality (B), and rather poorly for breakdowns (C). When there is private information, the Coase theorem predicts significantly worse in terms of efficiency, allocative neutrality, and breakdowns. We conclude that private information has a major impact in diminishing the predictive validity of the Coase theorem. In comparison, the generalized Myerson-Satterthwaite theorem predicts substantially better than the Coase theorem when there is private information in terms of inefficiency, allocative neutrality, and frequency of bargaining breakdown.

References

Coase, Ronald (1960), "The Problem of Social Cost," *Journal of Law and Economics,* 3, 1-31.

Coase, Ronald (1990), *The Firm, the Market, and the Law,* (Chicago: University of Chicago Press).

Demsetz, Harold (1966), "Some Aspects of Property Rights," *Journal of Law and Economics,* October, 61-70.

Fudenberg, D., and J. Tirole (1991), *Game Theory,* (Cambridge: MIT Press).

Hoffman, E. and M. L. Spitzer (1985), "The Coase theorem: Some experimental tests," *The Journal of Law and Economics*, 25, 73-98.

Holt, C. A., and D.D. Davis (1993), *Experimental Economics,* (Princeton: Princeton University Press).

Hurwicz, Leonid (1995), "What is the Coase Theorem?" *Japan and the World Economy,* 7, 49-74.

McKelvey, Richard and Talbot Page (1996), "The Coase Theorem with Private Information," mimeo.

Mishan, E. J. (1971), "The Postwar Literature on Externalities: An Interpretive Essay," *Journal of Economic Literature,*1-28.

Myerson, R. B., and M. A. Satterthwaite (1983), "Efficient mechanisms for bilateral trading," *Journal of Economic Theory*, 29, 265-281.

Ordeshook, P. (1986), *Game Theory and Political Theory,* (New York: Cambridge University Press).

Roth, A., and M. Malouf (1979), "Game theoretic models and the role of information in bargaining," *Psychological Review*, 574-594.

Samuelson, William (1984), "A Comment on the Coase Theorem" in *Game-Theoretic Models of Bargaining,* A. Roth, editor, (New York: Cambridge University).

Turvey, Ralph (1963), "On Divergences Between Social Cost and Private Cost," *Economica,* 309-13.